Renewable and Sustainable Energy Reviews 37 (2014) 123–141

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

# A review of methods to match building energy simulation models to measured data

Daniel Coakley [a,b,*], Paul Raftery [c], Marcus Keane [a,b]

[a] Department of Civil Engineering, NUI Galway, Ireland
[b] Informatics Research Unit for Sustainable Engineering (IRUSE), NUI Galway, Ireland
[c] Centre for the Built Environment (CBE), University of California, Berkeley, United States

ARTICLE INFO

*Article history:*
Received 19 March 2014
Accepted 3 May 2014
Available online 27 May 2014

*Keywords:*
Review
Calibration
Optimisation
Simulation
EnergyPlus
Uncertainty

ABSTRACT

Whole building energy simulation (BES) models play a significant role in the design and optimisation of buildings. Simulation models may be used to compare the cost-effectiveness of energy-conservation measures (ECMs) in the design stage as well as assessing various performance optimisation measures during the operational stage. However, due to the complexity of the built environment and prevalence of large numbers of independent interacting variables, it is difficult to achieve an accurate representation of real-world building operation. Therefore, by reconciling model outputs with measured data, we can achieve more accurate and reliable results. This reconciliation of model outputs with measured data is known as calibration.

This paper presents a detailed review of current approaches to model development and calibration, highlighting the importance of uncertainty in the calibration process. This is accompanied by a detailed assessment of the various analytical and mathematical/statistical tools employed by practitioners to date, as well as a discussion on both the problems and the merits of the presented approaches.



**Contents**



* Corresponding author at: Department of Civil Engineering, NUI Galway, Ireland. Tel.: +353 87 2285147.
*E-mail address:* daniel.coakley@nuigalway.ie (D. Coakley).

http://dx.doi.org/10.1016/j.rser.2014.05.007
1364-0321/

## 1. Introduction

In order to understand building energy simulation, it is necessary to understand scientific models in general. According to Saltelli et al. [1], models can be

- *Diagnostic* or *Prognostic*: Diagnostic models are used to identify the nature or cause of some phenomenon. In other words, it may be used to better understand the laws which govern a given system. Prognostic models, on the other hand, are used to predict the behaviour of a system, given a set of well-defined laws governing that system.
- *Law-Driven or Data-Driven:* Law-Driven (or forward) models apply a given set of laws (e.g., gravity, heat/mass transfer etc.) which govern a system, in order to predict its behaviour given system properties and conditions. Data-Driven (or inverse) models work on the opposite approach, using system behaviour as a predictor for system properties. Therefore, data-driven models can be used to describe a system with a minimal set of adjustable inputs [2]. In contrast, law driven models are often over-parameterised, in that they require more inputs than available data can support. However, the advantage of law-driven models is that they offer the ability to model system behaviour given a set of previously unobserved conditions, while data-driven models would require prior data in order to model behaviour. A simplified comparison of law-driven and data driven models is presented in Fig. 1 [3].

Building energy simulation (BES) models, as used in building design, can generally be classified as prognostic law-driven models in that they are used to predict the behaviour of a complex system given a set of well-defined laws (e.g., energy balance, mass balance, conductivity, heat transfer, etc.).

Conversely, data-driven (inverse) approaches, in the context of building energy modelling, refer to methods which use monitored data from the building to produce models which are capable of accurately predicting system behaviour. Inverse methods for energy use estimation in buildings can be broadly classified into three main approaches [4] (Table 1),

(i) Black-box approach: This refers to the use of simple mathematical or statistical models (e.g., regression, neural-networks etc.) which relate a set of influential input parameters (e.g., occupancy and weather) to measured outputs. Model input coefficients are determined such that they produce an algorithm with the ability to predict system behaviour. It is important to note that these input coefficients have no direct link to a definitive parameter in the physical environment.
(ii) Grey-box/parameter estimation: Grey box approaches differ from black-box approaches in that they use certain key (or aggregated) system parameters identified from a physical system model.
(iii) Detailed model calibration: The final approach uses a fully-descriptive law-driven model of a building system and tunes the various inputs to match the measured data.This approach provides the most detailed prediction of building performance, given the availability of high-quality input data. Since it is explicitly linked to physical building, system and environmental paramters, it provides a platform for assessing the impact of changes to these parameters (e.g., retrofit analysis).

### 1.1. Building energy performance simulation (BEPS) tools

Whole building energy simulation tools allow the detailed calculation of the energy required to maintain specified building performance criteria (e.g., space temperature and humidity), under the influence of external inputs such as weather, occupancy and infiltration. Detailed heat-balance calculations are carried out at discrete time-steps based on the physical properties of the building and mechanical systems as well as these dynamic external inputs (weather, occupancy, lighting, equipment loads etc.). These calculations are generally performed over the course of a full year. These tools fall into the category of prognostic law-driven simulation tools. Some of the main tools which will be discussed during the course of this review are

- DOE-2 [5] is a freeware building energy simulation tool which predicts the hourly energy use and energy cost of a building given hourly weather information, a building geometric and HVAC description, and utility rate structure. Its development was funded by the U.S. Department of Energy (DOE), hence the name.
- EnergyPlus [6] is an advanced whole building energy simulation tool, developed on the basis of work carried out on DOE-2. It incorporates the same functionality as DOE-2, producing hourly (or sub-hourly) energy costs of a building given system input information. It also incorporates many advanced features not available in DOE-2, such as multi-zone airflow and extensive HVAC specification capabilities.
- TRNSYS [7] is a transient system simulation program with a modular structure which implements a component-based simulation approach. Components may be simple systems like pumps or fans, or complex systems such as multi-zone buildings.
- ESP-r [8] is an integrated modelling tool for the simulation of the thermal, visual and acoustic performance of buildings. Similar to EnergyPlus and DOE-2, ESP-r requires user-specified information regarding building geometry, HVAC systems, components and schedules. It supports explicit energy balance in each zone and at each surface as well as incorporating inherent uncertainty and sensitivity analysis capabilities.

The above four simulation programs represent the most common tools encountered in conducting this review. However, many more tools are available, some of which are tailored specifically to certain tasks (e.g., HVAC simulation, solar gain, daylighting etc.). Crawley et al. [9] presents a comparison of the main features and capabilities of the top 20 tools available at the time of publication.

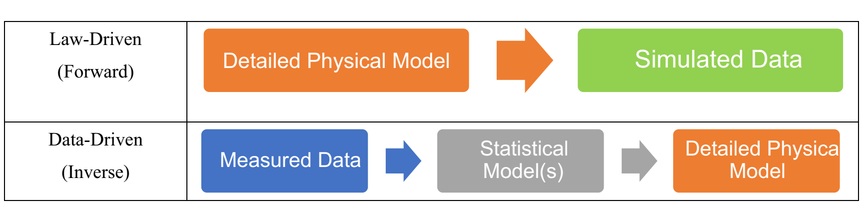



**Fig. 1.** Law-driven (forward) models vs. data-driven (inverse) models.

**Table 1**
Comparison of approaches to building energy simulation.

| Approaches | Advantages | Disadvantages |
|---|---|---|
| **Black-box** | • Very short development time.<br>• Provides an accurate predictor of building performance, given quality prior training data. | • Requires extensive training data for performance prediction.<br>• No explicit link between model inputs and physical building parameters – impossible to extrapolate model to compute effect of design or operational changes.<br>• Requires re-training when changes are made to building fabric, schedules or operation. |
| **Grey-box** | • Shortened development time by combining engineering model with statistical models.<br>• Accurate predictor of building performance, given quality prior training data.<br>• Linked to aggregated physical building, system and environmental parameters. | • Requires re-training when changes are made to building fabric, schedules or operation.<br>• Requires high level of knowedge of both engineering models and statistical models for development.<br>• Can only be extrapolated to account for chances to aggregate/simplified input parameters. |
| **Detailed simulation** | • Provide a detailed prediction for building energy performance.<br>• Linked to specific physical building, system and environmental parameters. | • Over-parameterised and under-determined.<br>• Require significant time, effort and expertise for development. |

**Table 2**
Model development and calibration issues.

| BEPS modelling issues | BEPS calibration issues |
|---|---|
| • **Standards**: Lack of understanding and consistent use of standardized methods.<br>• **Expense**: The time, knowledge, expertise and cost required to develop accurate models of building geometry and HVAC systems.<br>• **Integration**: Poor integration between various 3D modelling software packages (such as Autodesk Revit and ArchiCAD) and BEPS simulation packages (such as EnergyPlus, TRNSYS and Modelica). | • **Standards**: Lack of explicit standards for calibration criteria – current guidelines only specify acceptable error ranges for yearly whole-building simulation, but do not account for input uncertainty, sub-metering calibration, or zone-level environmental discrepancies.<br>• **Expense**: The expense and time needed to obtain the required hourly sub-metered data, which is usually not available.<br>• **Simplification**: Calibration is an over-specified and under-determined problem. There are thousands of model inputs but relatively few measurable outputs with which to assess the model accuracy.<br>• **Inputs**: Lack of high-quality input data required for detailed models.<br>• **Uncertainty**: There are currently few studies which account for uncertainty in model inputs and predictions, thus leading to a lack of confidence in BES outputs.<br>• **Identification**: Problems identifying the underlying causes of discrepancies been model predications and measured data.<br>• **Automation**: Lack of integrated tools and automated methods that could assist calibration. |

### *1.2. Benefits of BEPS*

While the initial focus of BEPS tools was primarily on the design phase, simulation is now becoming increasingly relevant in post-construction phases of the building life-cycle (BLC), such as commissioning and operational management and control [10]. Since BEPS models are based on physical reality rather than arbitrary mathematical or statistical formulations, they have a number of inherent advantages. One of the primary benefits of detailed simulation models over statistical models is their ability to *predict system behaviour given previously unobserved conditions*. This allows for analysts to make alterations to the building design or operation while simultaneously monitoring the impact on system behaviour and performance. Despite the potential benefits and the significant progress which has been made in the development of advanced simulation programmes capable of modelling complex systems and environments, there still remain a number of problems which inhibit their widespread adoption

## 2. Problems with BEPS and model calibration

At present, building energy performance simulation models (BEPS) are under-utilised within the AEC industry for a number of reasons, some of which were highlighted in a recent Rocky Mountain Institute (RMI) study [11]. These can be broadly grouped into two main categories, modelling and calibration, as described in Table 2.

Numerous studies [12–14] have indicated discrepancies, often significant (up to 100% differences), between BEPS model-predicted and the actual metered building energy use. This undermines confidence in model predictions and curtails adoption of building energy performance tools during design, commissioning and operation. In order for BEPS models to be used with any degree of confidence, it is necessary that the existing model closely represent the actual behaviour of the building under study. This can be achieved through model calibration, the purpose of which is to reduce the discrepancies between BEPS prediction and measured building performance. However, the calibration of forward building energy performance simulation (BEPS) programs, involving thousands of input parameters, to commonly available building energy data is a highly under-determined problem which yields multiple non-unique solution [15,16]. As a result, calibration methodologies and results are often not discussed in detail in many case studies. An approach in which the analyst tunes, or “fudges” [17], some of the myriad of input parameters until the model meets the acceptance criteria is commonly used. This is not conducive to the development of reliable building energy

simulation models. The current approaches to model calibration and their limitations are discussed in further detail in Sections 5 and 6.

## 3. Methods for assessing calibration performance

In the early years of building simulation, simple per cent difference calculations were the primary means of comparing measured and simulated data [18–20]. However, as noted by Diamond and Hunn [18] this often led to a compensation effect, whereby over-estimations cancelled out under-estimations. Bou-Saada and Haberl [21] proposed the adoption of standardised statistical indices which better represent the performance of a model [21–23]

- Mean Bias Error (MBE) (%): This is a non-dimensional bias measure (i.e., sum of errors), between measured and simulated data for each hour. The MBE is a good indicator of the overall bias in the model. It captures the mean difference between measured and simulated data points. However, positive bias compensates for negative bias (the cancellation effect). Hence, a further measure of model error is also required.

$$\mathrm{MBE}\ (\%) = \frac{\sum_{i=1}^{N_p}(m_i - s_i)}{\sum_{i=1}^{N_p}(m_i)} \tag{1.1}$$

  where $m_i$ and $s_i$ are the respective measured and simulated data points for each model instance '$i$' and $N_p$ is the number of data points at interval '$p$' (i.e., $N_{monthly}=12$, $N_{hourly}=8760$).
- Root Mean Square Error (RMSE) (%): The root mean square error is a measure of the variability of the data. For every hour, the error, or difference in paired data points is calculated and squared. The sum of squares errors (SSE) are then added for each month and for the total periods and divided by their respective number of points yielding the mean squared error (MSE); whether for each month or the total period. A square root of the result is then reported as the root mean squared error (RMSE).
- Coefficient of Variation of Root Mean Square Error CV(RMSE) (%):This index allows one to determine how well a model fits the data by capturing offsetting errors between measured and simulated data. It does not suffer from the cancellation effect.

$$\mathrm{CV\ RMSE}\ (\%) = \frac{\sqrt{\left(\sum_{i=1}^{N_p}(m_i - s_i)^2 / N_p\right)}}{\overline{m}} \tag{1.2}$$

where $m_i$ and $s_i$ are the respective measured and simulated data points for each model instance '$i$'; $N_p$ is the number of data points at interval '$p$' (i.e., $N_{monthly}=12$, $N_{hourly}=8760$) and $\overline{m}$ is the average of the measured data points.

The validation of building energy simulation models is currently based on a models compliance with standard criteria for CVRMSE and MBE, as shown in Table 3. These criteria vary depending on whether models are calibrated to monthly or hourly measured data, and are based on standard statistical indices.

Currently, building energy simulation models are generally considered 'calibrated' if they meet the criteria set out by ASHRAE Guideline 14 [24]. This means that once there is reasonable agreement between measured and simulated data, the model may be deemed 'calibrated' according to current international acceptance criteria for BEPS models. However, the model that meets these criteria is not unique and thus there are numerous models of the same building that can be considered to be 'calibrated'. In addition, it should be noted that current calibration criteria relate solely to predicted energy consumption, and do not account for uncertainty or inaccuracies of input parameters, or the accuracy of the simulated environment (e.g., temperature profiles).

**Table 3**
Acceptance criteria for calibration of BEPS models.

| **Standard/guideline** | Monthly criteria (%) | | Hourly criteria (%) | |
|---|---|---|---|---|
| | MBE | CVRMSE (monthly) | MBE | CVRMSE (hourly) |
| ASHRAE Guideline 14 [24] | 5 | 15 | 10 | 30 |
| IPMVP [25] | 20 | – | 5 | 20 |
| FEMP [26] | 5 | 15 | 10 | 30 |

## 4. Uncertainty in building simulation

In order to holistically address the topic of model calibration it is important to also consider the issue of *model uncertainty*, particularly for indeterminate models of complex systems. This is an important issue which is often neglected in BEPS calibration studies published to date and is not accounted for by any means in the current BEPS validation criteria. Models of complex systems are notoriously difficult to validate and have been the subject of much scientific discussion and debate in terms of quality and uncertainty [27]. Much of the reason for this debate stems from the fact that models of complex systems represent essential simplifications and simulation constraints. In other words, "the portion of the world captured by the model is an arbitrary enclosure of an otherwise open, interconnected system" [28]. This is particularly true when the purpose of the model is to provide some insight into the non-observable parts of the system. Thus mathematical formalisations of partially-observed experiments, even for well-defined or closed systems, can generate non-equivalent descriptions of these system (i.e., models whose outputs are compatible with the same set of observations but whose structures are not reconcilable with one another) [1]. This has also been referred to as equifinality [29,30] or model indeterminacy [1,31].

The built environment in particular presents a complex challenge in terms of energy modelling and accurate prediction. Any given building is characterised by a multiplicity of parameters including materials properties, occupancy levels, equipment schedules, HVAC and plant operation, climate and weather. These represent diverse sources of model parameter uncertainty. However, this does not illustrate the entire range of potential uncertainty encapsulated by any given building model. Numerous studies have focused on this problem [32–37], although few published case studies incorporate this work into their analyses. De Wit [32] classified the various sources of uncertainty in building performance simulation as follows:

- **Specification uncertainty**: arising from incomplete or inaccurate specification of the building or systems modelled. This may include any exposed model parameters such as; geometry, material properties, HVAC specifications, plant and system schedules etc.
- **Modelling uncertainty:** simplifications and assumptions of complex physical processes. These assumptions may be explicit to the modeller (zoning and stochastic process scheduling) or hidden by the tool (calculation algorithms).
- **Numerical uncertainty:** errors introduced in the discretisation and simulation of the model.
- **Scenario uncertainty:** external conditions imposed on the building, including outdoor climate conditions and occupant behaviour.

It is important that these sources of uncertainty are identified and quantified when assessing model predicted performance. This is particularly important given the 'equifinality' of simulation models (i.e., multiple disparate models may provide the same

results). Depending on the application of the BEPS model, it is important to know the degree of uncertainty associated with particular elements of the model or underlying mathematical formulation. This paper deals primarily with 'specification' and 'modelling' uncertainty and how this can be systematically propagated throughout the simulation model development process.

## 5. Current approaches to BEPS calibration:

The main approaches to building energy performance simulation (BEPS) model calibration were first classified by Clarke et al. [38] and adopted in a later literature review of calibration programs, tools and techniques by Reddy [39]. The four classes proposed are

(i) Calibration based on manual, iterative and pragmatic intervention.
(ii) Calibration based on a suite of informative graphical comparative displays.
(iii) Calibration based on special tests and analytical procedures.
(iv) Analytical/mathematical methods of calibration.

These classifications have been further extended in this review. In general, it was found that approaches to the tuning of simulation models to measured data can be more broadly defined as either *manual* or *automated*.

(1) **Manual** – these approaches predominantly rely on iterative pragmatic intervention by the modeller. These include any methods which employs no form of automated calibration through mathematical/statistical methods or otherwise.
(2) **Automated** – automated approaches may be described as having some form of automated (i.e., not user driven) process to assist or complete model calibration.

Both manual and automated approaches may employ specific *analytical tools or techniques* to assist in the calibration process, while automated approaches employ *mathematical and statistical techniques* to reach their goal.

### 5.1. Analytical tools and techniques

These can be broadly classified as manual user-driven techniques, but may also be employed as part of an automated calibration process. A list of the main calibration tools and techniques has been compiled following an extensive review of methodologies and applications over the past three decades. For clarity, these are divided into four main categories and presented alongside the relevant key publications in Table 4.

- Characterisation Techniques: techniques based on the characterisation of the physical and operational characteristics of the building being modelled.
- Advanced Graphical Methods: the use of graphical representations of building data or statistical indices.
- Model Simplification Techniques: techniques which aim to reduce the complexity of simulation models by reducing or aggregating the number of simulation variables.
- Procedural Extensions: the use of standard processes or techniques to improve the simulation and/or model calibration process.

An exhaustive list of the papers mentioned in Table 5 is available at the end of this paper.

### 5.2. Mathematical/statistical techniques

Modern mathematical and statistical methods are increasingly being employed to assist the calibration process. Applications which employ one or more of these techniques at any stage in the process, have been classified as automated approaches within this framework. Some of the mathematical/statistical approaches employed in calibration studies to date are summarised in Table 5, under the following two main categories:

- Optimisation Techniques: This covers the general methods used to optimise prediction performance of any type of model.
- Alternative Modelling Techniques: This section covers alternatives to detailed model calibration, described as black-box or grey-box approaches.

### 5.3. Summary of manual calibration developments

Over the past three decades, many procedures have been proposed for the calibration of whole building energy performance simulation models. This section examines manual calibration procedures, chronologically highlighting the new techniques employed by various authors as well as where these techniques have been adapted and advanced.

#### 5.3.1. Characterisation techniques

Waltz [41] claims that the single most important factor in developing accurate computer models of existing buildings is developing an intimate knowledge of the physical and operational characteristics of the building being modelled. This section of the review covers techniques which have been used to develop an understanding of these characteristics.

*5.3.1.1. Building and site audits.* An energy audit can be defined as a process to evaluate where a building or plant uses energy, and identifies opportunities to reduce consumption. There is an existing consensus on the definition of three typical levels of building audit [90]

- **Level 1 – walkthrough**: This generally implies a tour of the facility and visual inspection of energy using systems. This also includes an evaluation of energy consumption data to analyse energy use quantities and pattern, as well as providing comparisons to industry averages or benchmarks.
- **Level 2 – standard audit**: Energy uses and losses are quantified through a more detailed review and analysis of equipment, systems and operational characteristics. On-site measurements may be used to quantify and assess efficiency of energy end-users. This audit also includes an economic analysis of energy conservation measures (ECM's).
- **Level 3 – investment grade**: This includes a more detailed review of energy use by function as well as a comprehensive evaluation of energy-use patterns. Energy simulation software is employed to predict year-round energy use, accounting for weather and system variables. The method also accounts for system interactions to prevent over-estimation of savings.

A summary of the main deliverables for the three levels of energy audits is presented in Fig. 2 [91].

Lyberg [92] provides a comprehensive handbook on energy auditing procedures, defining the auditing process as "a series of actions, aiming at breaking down into component parts and quantifying the energy used in a building, analysing the applicability, cost and value of measures to reduce energy consumption,

**Table 4**
Analytical tools and techniques.

| Acronym | Name | Description | Key Papers |
|---|---|---|---|
| **Characterisation techniques** | | | |
| **AUDIT** | Detailed audit | Detailed audits are often conducted prior to building model development in order to gain a better knowledge of the building systems and characteristics (Geometry, HVAC systems, Lighting, Equipment, and Occupancy Schedules). | [40,41] |
| **EXPERT** | Expert knowledge/ templates/model database | Approaches which utilise<br>• Expert knowledge or judgement as a key element of the process.<br>• Prior definition of typical building templates.<br>• Database of typical building parameters and components in order to reduce the requirement for user inputs during model development. | [42–45] |
| **INT** | Intrusive testing | Intrusive techniques require some intervention in the operation of the actual building, such as 'Blink Tests' whereby groups of end-use loads (e.g., plugs loads, lighting etc.) are turned on and off in a controlled sequence in order to determine their overall impact on the baseline building load. | [46] |
| **HIGH** | High-res data | Data is recorded at hourly (or sub-hourly) levels as opposed to utilising daily load profiles or monthly utility bill data. | [38,47–49] |
| **STEM** | Short-term energy monitoring | Metering equipment is used to record on-site measurements for a short period of time ( > 2 weeks). This may be used in identifying typical energy end-use profiles and/or base-loads. | [50–52] |
| **Advanced graphical methods** | | | |
| **3D** | 3D-graphical comparison techniques | Three-dimensional graphs are used to aid comparison and/or calibration of measured and simulated data. This technique allows users to visualise large quantities of data, compared to traditional 2-D scatter plots etc. which are overwhelmed when analysing large quantities of data points. | [47] |
| **SIG** | Signature analysis methods | Signature analysis techniques are a specific type of graphical analysis technique, typically used by HVAC simulation engineers to identify faulty parameters in Air-Handling Unit (AHU) simulation. They may also be used to develop optimised operation and control schedules. Signature analysis methods are commonly used for the calibration of models based on the simplified energy analysis procedure (SEAP). | [53,54] |
| **STAT** | Statistical displays | This refers to the graphical representation of statistical indices and comparisons for easier interpretation. This includes data comparison techniques such as carpet plots, box-whisker mean (BWM) plots and monthly per cent difference time-series graphs. | [47] |
| **Model simplification techniques** | | | |
| **BASE** | Base-case modelling | The base-case model refers to the use of measured base-loads to calibrate the building model. Base-loads refer to minimum, or weather independent, electrical and gas energy consumption. Calibration is carried out during the base-case when heating and cooling loads are minimal and the building is dominated by internal loads, thus minimising impact of weather dependent variables. | [55,56] |
| **MPE** | Model parameter estimation | Deduction of overall aggregate (or lumped) parameters (such as *U*-values) using non-intrusive measured data. | [57] |
| **PARRED** | Parameter reduction | This involves reducing the requirement for detailed input for variable schedules (e.g., plug loads, lighting, occupancy, equipment etc.). Day-Typing is one such approach which works by analysing long-term data and reducing this to manageable typical day-type schedules (e.g., weekday's vs. weekends, winter vs. summer). Zone-typing may also be used to reduce large models into similar thermal zones (e.g., Core, Perimeter, Offices, Unoccupied spaces etc.). | [19,20,58,59] |
| **DISSAG** | Data disaggregation | Data disaggregation refers to the application of non-intrusive techniques to de-couple multiple measured data streams (e.g., energy end-use data from whole-building electrical energy consumption). | [60–62] |
| **Procedural extensions** | | | |
| **EVIDENCE** | Evidence-based model development | For the purpose of this review, evidence-based approaches may be described as those that implement a procedural approach to model development, making changes according to source evidence rather than ad-hoc intervention. Strictly, this approach should account for adjustments to model parameters in a structured fashion (e.g., using version control software). | [21,48,59,63] |
| **SA** | Sensitivity analysis | Sensitivity analysis procedures may be employed in some studies to assess the influence of input parameters on model predictions. This information may be used to identify important parameters for measurement or detailed investigation. | [64–68] |
| **UQ** | Uncertainty quantification | This refers to assessment of parameter uncertainty as part of the calibration process. This information may be used to directly assist in model calibration or provide a basis for risk quantification within the results (e.g., uncertainty related risk quantification in ECM analysis). | [32,33,35,69,70] |

and recommending what measures to take". Lyberg proposes a staged audit process

(1) Building Rating – assessing potential high-potential buildings for audit.
(2) Disaggregation of energy consumption (Refer Section 5.3.2.5).
(3) Energy Conservation Opportunity (ECO) identification.
(4) ECO evaluation.
(5) Post Implementation Performance Analysis (PIPA).

An extensive collection of necessary audit templates are also provided in Volume II of the audit handbook [92], categorised under the following 4 headings: (1) Audit Procedures, (2) Measurement Techniques, (3) Analysis Techniques and (4) Reference Values.

Waltz [41] suggests two types of survey (1) observational; and (2) electrical load survey (see Section 5.3.1.2). The observational survey refers to the actual functioning of the buildings control systems as opposed to relying on documentation and as-built drawings. Often controls may not be installed as per the design documentation, or operational controls may have been overridden, or have simply failed. The authors also suggest a "late-night" tour of the facility and its HVAC systems to determine 'actual' operating schedules, which often differ from those prescribed in operation & maintenance (O&M) documentation.

CEC [93] provides a comprehensive guide for reporting investment grade audits of various types of facilities and project types (e.g., lighting, HVAC, and cogeneration). The guide also includes a copy of sample field data sheets for recording site specific

**Table 5**
Mathematical and statistical calibration techniques.

| Acronym | Name | Description | Key papers |
|---|---|---|---|
| **Optimisation techniques** | | | |
| **BAYES** | Bayesian calibration | Bayesian calibration is an alternative statistical approach to model calibration. The approach offers the advantage of naturally accounting for uncertainty in model prediction through the use of prior input distributions. | [71–74] |
| **OBJECT/ PENALTY** | Objective/penalty function | Most mathematical techniques employ some form of optimisation function to reduce the difference between measured and simulated data. An objective function may be used to set a target of minimising, for example, the mean square error between measure and simulated data. Conversely, a penalty function may also be employed to reduce the likelihood to deviating too far from the base-case. | [15,42,75] |
| **Alternative modelling techniques** | | | |
| **ANN** | Artificial neural networks | Neural networks are computational models consisting of an interconnected group of artificial neurons. They are used for modelling complex relationships between inputs and outputs or for finding patterns in data. | [76–79] |
| **PSTAR** | Primary and secondary term analysis and re-normalisation | Analytical tool for the meaningful estimation of parametrs of a complex building from a few data channels over a short period (a few days). An 'audit' description of the building (capturing nomincal building fabric parameters) is used to estimate heat-flows. These heat-flows are then re-normalised to saisfy an energy-balance equation using a least-squares method. | [52,80,81] |
| **META** | Meta modelling | The use of computationally efficient analytical surrogate models which emulate the performance prediction of their complex engineering-based counterparts. | [82,83] |
| **SEAP** | Simplified energy analysis procedure | The simplified energy analysis procedure refers to the use of simplified engineering models to represent the building. This may be accomplished by dramatically reducing the number of zones or AHU's in the model by grouping them together. | [84–86] |
| **SYS** | Systems identification | This technique refers to the process of constructing models based only on the observed behaviour of the system (outputs) and a set of external variables (inputs), instead of constructing a detailed model based on 'first principles' of well-known physical variables. | [87–89] |

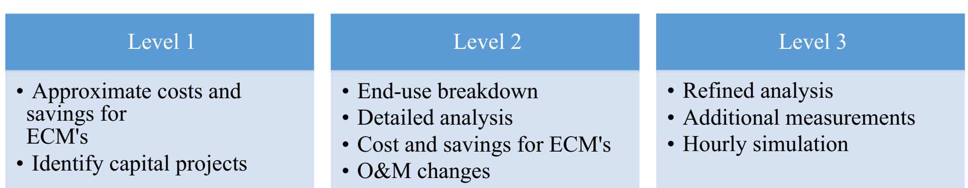



**Fig. 2.** Deliverables for energy audit levels 1, 2 and 3 [91].

information such as building data, occupancy schedules, lighting and equipment surveys as well as HVAC equipment data.

Ganji and Gilleland [94] provide assessment of investment grade energy audits and a review of typical cases, identifying several major shortcomings including lack of consistency in auditing, reporting and over-estimation of savings. These shortcomings stem from a number of deficiencies including a lack of expertise and fundamental engineering knowledge on the part of the surveyor. A lack of training in advanced energy simulation software was also identified as an issue, resulting in incorrect outputs in many cases.

Shapiro [40] also identifies shortcomings in the current approaches to commercial building audits, including a lack of clearly defined boundaries and limitations of simple building audits (Level 1 and Level 2). Shapiro proposes a comprehensive building audit on a room-by-room basis, capturing room-specific opportunities and documenting recommendations in the audit report. Improvements should focus not only on efficiency, but ensuring that the equipment meets the load requirements for the space. An example of a comprehensive lighting audit is given to illustrate how the proposed approach differs from standard walkthrough audits. The author identifies overlit areas and recommends multiple improvements (delamping, occupancy sensors and control changes). In contrast, a typical walkthrough audit would record existing equipment but may miss energy reduction recommendations. The proposed comprehensive audit approach is also applied to a case-study office building, identifying 46% potential energy savings, compared with 7% identified through a standard walkthrough audit.

To date, a number of standard auditing and energy assessment procedures have been proposed for different industries and applications [95]

- AuditAC: Developed as part of a European project "Field Benchmarking and market Development for Audit Methods in Air Conditioning". The project focused on providing tools and information for air-conditioning engineers to identify energy savings in HVAC systems [96].
- IEA Annex 11: Comprehensive handbook on energy auditing procedures developed in conjunction with the International Energy Agency (IEA) [92].
- AS/NZS 3598:2000: Standard developed by Australia and New Zealand energy authorities, targeting the commercial and industrial sector. The standard sets out minimum requirements

for commissioning and conducting energy audits which identify cost effective opportunities to improve efficiency and effectiveness in the use of energy [97].
- RP-351 Energy Audit Input Procedures and Forms: General ASHRAE procedures for energy auditing including an assessment of existing audit procedures [98].
- ASHRAE Procedures for Commercial Building Energy Audits: Standard for energy companies conducting energy audits of commercial buildings, including definitions of Levels 1, 2 and 3 audits [91].

*5.3.1.2. Short-term end-use monitoring (STEM).* STEM refers to the application of specialized software and hardware tools to systematically gather and analyse data typically over a short (typically two week) period to evaluate the performance of building energy systems, such as HVAC, controls, and lighting. Diagnostics based on short-term monitoring can clarify how the systems in a building actually perform, as well as highlighting key energy end-users.

A study by the Tishman Research Corporation [99] on the calibration of a DOE-2 office model to measured data was the first identifiable study which incorporated short-term end use monitoring to increase the accuracy of model inputs. Measurement errors for sensors were also accounted for in the study, showing an acceptance of potential uncertainties in the measured end-use values as opposed to solely model inputs.

Waltz [41] suggests measuring instantaneous power draw for every electrical panel or piece of equipment using a hand-held power factor metre. This is particularly important when high levels of accuracy are required, for example in high-rise multi-zone office buildings.

Kaplan et al. [16,19] suggest calibrating models to short typical periods as opposed to full year data, for example one month during a heating and cooling season. The authors incorporate short-term energy monitoring during these periods to assist calibration. Statistical analysis is applied to these short-term monitored end-uses to generate manageable DOE-2 schedules for lighting, equipment, occupancy setpoints etc. In this regards, monitored data is used to generate DOE-2 inputs and validate outputs.

A similar approach is adopted by Soebarto [46] for calibrating models to utility bill data using only two to four weeks measured data. The procedure requires the use of STEM in order to develop a set of energy end-use profiles, including; electrical energy, heat energy and, indoor temperatures. The author also proposes the use of intrusive blink-tests (see Section 5.3.1.4).

Short-term monitoring has since been used in a number of studies to assist in identifying input parameters [49,50].

*5.3.1.3. High-resolution data.* Clark et al. [38] investigated the use of calibrated ESP-r simulation to investigate the performance of passive solar components (PASSYS). The study was differentiated by its use of high-quality, high-resolution data and empirical evidence for model calibration and validation. First, a sensitivity analysis (SA) is carried out to quantify uncertainty bands associated with model predictions and associated parameter sensitivities. This information is used to design an experiment to capture a high-quality data set with which to quantify model residuals and identify their cause. The authors also highlight the importance of uncertainties when extrapolating from test-cell scenarios to full-scale application.

A study by Norford et al. [100] investigated the two-fold differences between a simulation model at design stage and actual operation for a low-energy office building. Focus was placed on high levels of instrumentation (100 sensors polled 200–300 times an hour) to provide hourly averages of ambient and interior conditions as well as energy consumption of HVAC and tenant equipment. The study concluded that differences were mainly due to unanticipated tenant energy consumption (64%), increased HVAC operation beyond design schedule (24%) and specification errors in HVAC equipment, building fabric and infiltration (12%). This highlights the importance of occupant behaviour in determining model performance as well as the need for sufficient instrumentation to monitor this behaviour if it is a significant factor in determining building performance.

*5.3.1.4. Intrusive testing.* An approach has been developed for determining characteristic building parameters using controlled heating and cooling tests over short periods of 3–5 days [51,52]. This test consists of a period of co-heating to determine an estimate for the building heat-loss co-efficient, and cool-down to provide an estimate for the effective thermal time constant of the building.

Soebarto [46] presents an approach for calibrating models to utility bill data using only two to four weeks measured data. A series of 'on–off tests' (or Blink Tests) were utilised to determine lighting and plug loads. In these tests, all electrical loads were turned off for a short period, and back on again. This equipment 'on–off' cycling is carried out in a predetermined pattern while recording electrical energy use on a data logger, in order to accurately determine the load profile for various equipment end-users without the need for individual sub-metering. This method resulted in an hourly calibration accuracy of 6.7% CV (RMSE) for whole building electricity and 1% for chilled water energy use.

*5.3.2. Advanced graphical approaches*

In the past, graphical techniques were confined to simple time-series plots [101]. With the increasing availability of detailed measured data and requirement to better understand this information, there has been extensive work carried out in the area of graphical data representations.

*5.3.2.1. 3-D comparative plots.* Bronson et al. [20] proposed a means of calibrating hourly building energy models to non-weather dependent (or scheduled) loads using novel comparative three-dimensional graphics which allowed hourly differences to be viewed for the entire simulation period. Day-typing was also used to assist in the calibration process. The authors reported that the availability of comparative three-dimensional surface plots significantly improved the ability to view small differences between the simulated and measured data, which allowed for the creation of a "super-tuned" DOE-2 simulation that matched the electricity use within 1%. The process of identifying and fixing unknown "misfits" between the simulation and the measured data was significantly enhanced by the use of the plots.

Bou-Saada and Haberl [21,47] propose the use of 3D surface plots and statistical indices (refer to Section 5.3.2.2) to provide a global view of the differences between measured and computed hourly values in order to help identify time-dependent patterns in discrepancies between measured and simulated data. McCray et al. [102] propose another graphical method to calibrate a DOE-2.1 model to one year of 15-min interval data for whole-building energy use. The Visual Data Analysis (VDA) method allows the modeller to quickly review the simulation results and make iterative changes to the models. A number of later studies focused on further developing this approach by means of visual comparative displays [20,21,47,103–106].

Christensen [107] originally proposed the use of colour contour plots (or Energy Maps, EMAPS) to help display hourly data from a commercial building. Haberl et al. [106] adopted this technique in

developing graphical comparative displays with time-sequenced contour plots. Raftery and Keane [108] proposed the use of carpet contour plots as a means of speeding up the identification of major discrepancies between modelled and simulated data as well as a useful tool for fault detection.

*5.3.2.2. Graphical statistical indices.* Graphical statistical indices refer to the graphical representation of statistical indices through the use of graphical techniques. One such approach is binned box-whisker mean plots [21,109] which display maximum, minimum, mean, median, 10th, 25th, 75th and 90th percentile points for each data bin given a period of data. These plots eliminate data overlap and allow for more informative statistical characterisation of the dense cloud of data points. The authors also proposed the use of temperature bin analysis, 24-h weather day-type analysis and 52-week bin analysis. Further examples can be found in a number of more recent case studies illustrating the importance of effectively conveying statistical information behind calibration studies [48,55,108,110,111].

*5.3.2.3. Signature analysis.* One of the major issues in tackling building energy calibration is the issue of accurately modelling heating and cooling energy consumption. Katipamula and Claridge [84] proposed an approach for developing simplified system models for retrofit analysis, based on the work of Knebel [85] on the simplified energy analysis procedure (SEAP) (see Section 5.4.2.4). This was later extended to account for calibration and development of optimised control strategies [86]. Based on this work, a process was developed to generate graphical signatures of heating and cooling energy consumption [112–114]. The authors proposed that these graphical signatures would allow simulation engineers to identify the impacts of different input parameters (weather, occupancy, outside air intake, system type etc.) on an AHU's heating and cooling energy consumption. In addition, the technique may be used by commissioning engineers to identify faulty parameters, and develop optimised operation and control schedules.

Liu et al. [53] propose a step-by-step procedure for the manual calibration of simulation models, based on the definition of two characteristic signatures

- Calibration Signature: normalised plot of the difference between measured energy consumption values and the corresponding simulated values as a function of outdoor air temperature. For a given system type and climate, the graph of this difference has a characteristic shape that depends on the reason for the difference.
- Characteristic Signature: By simulating the building with one value for an input parameter (the "baseline" run), then changing that input parameter by a given amount and rerunning the simulation, the "residuals" between these two simulations can be calculated, normalised, and plotted vs. outdoor air temperature, producing a characteristic signature. By matching the observed signature with the published characteristic signature, the analyst is given clues to the factors that may be contributing to the errors he or she is observing.

Liu and Liu [54] provide a rapid two-stage calibration procedure for simplified energy models, based on the use of calibration signatures. A simplified model of a high-rise office building is developed and calibrated to two weeks worth of measured data. This model is then used to simulate the hourly heating and cooling energy consumption for the building. Calibration signatures are then used to compare measured and simulated data in order to give an indication of which parameters should be changed and the corresponding magnitude of change required. A second stage of calibration requires the fine-tuning of these parameters to obtain a better overall fit of the model to measured data. The authors present a case study which serves to highlight a number of issues with the calibration process. Firstly, this type of parameter tuning is typical of the general approach to model calibration, and while it may serve to produce a model which demonstrates sufficient overall accuracy when compared to measured data, it is probably not a good representation of the actual building being analysed. It is also highly dependent on analyst knowledge and skill, data availability, and allowed time-frame. The authors also point out that the satisfaction of hourly ASHRAE calibration criteria is quite difficult, even when high levels of measured data are available. It is also questionable as to whether it is even useful (or appropriate) to fine-tune a model to a very high degree of accuracy when employing generalised model assumptions and typical operation profiles.

*5.3.2.4. Parameter reduction (day-typing and zone-typing).* The process of parameter reduction or simplification relies on the statistical characterisation of complex inputs in order to reduce the number of inputs in a model. One approach which has been used extensively is day-typing, in which building energy use is characterised on a daily profile, rather than on an hourly basis. This approach allows for the definition of typical days (e.g., weekdays, weekends, and holidays) which can be used to characterise building energy use, thus condensing a large quantity of complex measured building data into relatively few input points or schedules.

Kaplan et al. [16,19] use day-typing to group days with reasonable uniform non-HVAC load shapes. Zone-typing (i.e., grouping similar zones) is used to further apply these day-types across multiple zones. Bronson et al. [20] uses day-typing routines (for occupancy and equipment scheduling) to calibrate a DOE-2 simulation model. Hadley [58] uses a combination of principal component analysis and cluster analysis to identify distinctive weather day types (which represent repeatable weather conditions that typically occur at each site) from one year of National Weather Service (NWS) station data. HVAC system energy consumption data for each day are then grouped by these weather day types, and daily total and hourly load profiles were developed for each day type.

Raftery et al. [48,59] incorporate zone-typing to separate thermal zones in such a way as to minimise inaccuracies incurred by representing multiple actual thermal zones in a building with a single large zone in the model. This is achieved by assigning thermal zones in the model based on four major criteria, (1) space function, (2) position relative to exterior, (3) available measured data, and (4) space conditioning method.

*5.3.2.5. Data disaggregation.* Disaggregation is the splitting up of the total building energy consumption into its component parts. There are a number of reasons as to why this is done, i.e., to focus on specific energy flows and identify areas for retrofit and conservation. Lyberg [92] proposes data disaggregation as part of a staged audit process as a means of focusing attention on high-importance areas. This can help limit subsequent auditing to the areas where the most productive retrofits could be carried out. This step will directly assist in the identification of energy-conservation opportunities (ECO's).

Akbari [60,62] developed an algorithm to disaggregate short-interval (hourly) whole building electrical load into major end-uses. The end-use disaggregation (EDA) algorithm utilises statistical characteristics of measured hourly, whole-building load and its inferred dependence on temperature to produce hourly load

profiles for air-conditioning, lighting, fans, pumps and miscellaneous loads. Regression models are developed for each hour of the day for major day types (see Section 5.3.3.3) between measured building energy use and outdoor dry-bulb temperature. Since the temperature dependency of the building may change with season, the author suggests using two season specific (summer and winter) sets of temperature regression coefficients. The regression constant for these models are assumed to provide an indication of the weather-independent energy use, while the slope represents weather-dependent behaviour. Since the regression models provide no information about the breakdown of the temperature-independent load, it is simply pro-rated against loads predicted by simulation as well as on-site measurements. The approach is applied to numerous retail and commercial facilities [60–62]. The authors conclude that this is a useful approach for buildings in which the whole-building temperature dependent load is primarily due to the HVAC system (i.e., only the HVAC load is sensitive to outdoor temperature). This assumption may be applied to large offices and commercial buildings, but not to buildings characterised by non-HVAC end-uses such as refrigeration (which is weather dependent).

### 5.3.3. Procedural extensions

The following section describes procedural tools and techniques used to assist in improving the overall calibration process.

*5.3.3.1. Evidence-based development.* Manual approaches to model calibration generally rely on manual pragmatic user intervention to 'fine-tune' individual parameters to achieve a calibrated solution. However, these changes are often not tracked or recorded, and are rarely reported. This results in a situation whereby the calibration process relies heavily on user knowledge, past experience, statistical expertise, engineering judgement, and an abundance of trial and error [21]. In order to improve the reliability and reproducibility of the calibration process it is necessary to keep a history of the decisions made along with the evidence on which these decisions were based [21,59]. This allows future users to review the entire calibration process and the evidence on which the model is based. In addition, changes to the input parameters should only be made according to available evidence and clearly defined priorities [59]. A number of studies incorporate systematic evidence-based model development at the core of the calibration process [21,49,55,56,63,115,116].

*5.3.3.2. Sensitivity analysis (SA).* Sensitivity analysis has been employed in recent calibration efforts to identify parameters of greatest influence on energy end-use in a building. There are a number of available techniques available for conducting sensitivity analyses, depending on the particular requirements and application (e.g., single vs. multiple parameters). For detailed descriptions of tools and techniques, refer in particular to the work of Saltelli et al. [117,118] on this particular subject.

Clarke et al. [38] used two sensitivity analysis techniques to determine uncertainty bands associated with ESP-r predictions. Differential sensitivity analysis (DSA) was used to determine total uncertainty band as the root mean squared summation of individual uncertainties due to each input parameter. Monte-Carlo sensitivity analysis (MCSA) was used to determine the total uncertainty band by perturbing all the input parameters simultaneously. These sensitivity methods have been incorporated into ESP-r simulation software [8,35] for the purpose of uncertainty analysis.

Westphal and Lamberts [68] present a calibration study of a 26,264 $m^2$ public office building, combining a building energy audit, model sensitivity analysis and manual tuning of influential parameters. The study concludes with an electricity consumption prediction within 1% of the measured values within four iterations of the base case model.

*5.3.3.3. Uncertainty quantification.* As identified by Carroll and Hitchcock [15], there exist multiple solutions which produce good overall agreement with measured data even though individual parameters are incorrectly defined. Hence, if using these inputs to infer any sort of meaning (e.g., for ECM analysis), it is important to account for uncertainty in these inputs. Reddy [39], states that uncertainties in building simulation generally arise from four main sources

(i) Improper input parameters.
(ii) Improper model assumptions.
(iii) Lack of robust and accurate numerical algorithms.
(iv) Error in writing simulation code.

While sources (ii)–(iv) deal directly with the simulation program and internal algorithms and assumptions, source (i) depends on the accuracy (and uncertainty) of the available input information. Since the validation of model algorithms is covered extensively in other studies [119–121], this review will focus on contributions to the identification of error and uncertainty in model input parameters, and how this has been applied to model calibration.

As discussed in the previous section, Clarke et al. [38] used sensitivity analysis to determine uncertainty bands associated with ESP-r predictions of internal air-temperature in his PASSYS test-cell experiments. In this case, uncertainty bands were quite narrow, reflecting the level of control of the experiments in terms of ESP-r input parameters. It was shown, however, that uncertainty bands were largely temperature-dependent, due primarily to the uncertainty in conservatory air temperature prediction. This was due to instrument accuracy for solar radiation measurement (varying by as much as $\pm 3\%$).

Lomas and Eppel [64] discuss the application of three sensitivity analysis techniques (DSA, MCSA, and SSA) to determining the relative sensitivities, in both hourly and daily average model predictions (using ESP-r, HTB2 and SERI-RES), due to the uncertainties in over 70 input parameters. Lomas et al. [122] conducted an extensive review of dynamic thermal simulation programs (DSPs) comparing measurements with predictions and accounting for experimental uncertainty. The authors state that total model uncertainty has two components: (1) measurement errors, as above – which are easy to identify; and (2) uncertainties in program input data – which is more difficult to calculate. This difficulty is due to the large number of inputs which require quantification of associated uncertainty, as well as the propagation of this uncertainty through the DSP to determine the overall prediction uncertainty.

De Wit and Augenbroe [32] address uncertainties in building performance evaluations and their potential on design decisions. The authors examine uncertainties in material properties as well as those stemming from model simplifications. They suggest a statistical screening technique (using Monte-Carlo analysis) to determine which sources have dominant effects on the outcome of the simulation. The procedure is illustrated for a simple building envelope and considered parameters such as wind speed, indoor air distribution, and envelope material and heat transfer coefficients.

Reddy et al. [42,43] identified the necessity for uncertainty analysis, which had been over-looked in many calibration studies, particularly in ECM analysis applications. In this work, uncertainty is addressed by assigning ranges of variation to influential input parameters and a Latin-hypercube Monte-Carlo (LHMC) simulation is

carried out to produce multiple possible solutions. The author selects the top 20 solutions, rather than selecting a single solution, to produce a range of values for the predicted performance of ECM's (rather than a single value). Overall, the authors found the relative uncertainty (or fractional difference) between actual and predicted values to be in the range of 25–50%. However, in most cases, the actual savings are usually contained in the range predicted. In conclusion, the authors suggest that one should not rely on calibrated simulations which predict savings of less than 10% (as associated uncertainty could account for up to 50% of this value).

### 5.4. Summary of automated calibration developments

The following summarises the major developments in automated calibration of building energy performance simulation models over the last three decades.

#### 5.4.1. Optimisation techniques

*5.4.1.1. Objective function.* The first automated calibration technique called renewable energy savings estimation method (RESEM) was used for evaluation of ECM's using pre-retrofit and post-retrofit data [45,123,124]. The tool is based on a previously developed set of *knowledge-based expert rules* designed to bridge simulation models with measured utility bill data [44]. Retrofit energy savings estimation model (RESEM) uses a self-contained energy simulation programme similar to DOE-2, called RESegy. The goal of the project was to provide a simple cost-effective solution for ECM analysis by staff with little or no energy simulation expertise. As such, it relied on a database of expert knowledge for the development of building prototypes and parameter defaults based on minimal information from the user. The tool was benchmarked against DOE-2 using a simple base-case building. Comparisons of monthly heating and cooling loads (including peak loads) as well as electrical and gas energy consumption, as computed by DOE-2.1E and RESEM, were performed.

Lavigne [125] implemented a similar DOE-2 based assisted calibration process using built-in engineering rules as well as optimisation algorithms based on a Maquardt–Levenberg non-linear least squares method. Two real case studies are presented and calibrated to monthly utility bill data by tuning a set of user-defined parameters until acceptable limits are reached. In the presented case studies, this was achieved in 2–3 iterations, achieving a monthly and annual difference in measured and simulated energy consumption of 10.9% and −1.1% respectively.

*5.4.1.2. Penalty function.* Based on their original experience with RESEM, Carroll and Hitchcock [15] introduced a more generic approach to systematically adjusting ("tuning") the parameters of a simulatable building description in order to match simulated performance to metered utility data. The underlying method is based on the minimisation of differential terms between measured and simulated data. In addition, the approach incorporates a weighting function to describe the relative importance of any single term within the minimisation function, thus maintaining reasonable parameter values during the calibration process. The paper also addresses two other important issues

- Existence – it may not be possible to find an exact match between measured and simulated performance (i.e., the simulation model does not represent exactly what happens in the real building). Therefore, rather than identify an exact solution, the authors suggest finding a minimum quantity based on the normalised difference between predicted and actual consumption.
- Uniqueness – there may be many solutions which match the defined minimisation criteria. This can be addressed by providing additional matching constraints in the minimisation function, thus reducing the number of possible solutions. The authors suggest the use of a penalty function term which increases quadratically with the difference between each input parameter and its corresponding preferred value.

The approach utilises a prototype building generator to assist in the creation of the initial building model. The tuning process relies on some knowledge of the building to decide on parameters for adjustment, based on associated uncertainties classified during a building audit.

More recently, a methodology has been developed for the systematic calibration of energy models that includes both parameter estimation and determination of uncertainty in the calibration simulation [42,75]. Based on the building type, the user must heuristically define a set of influential parameters and schedules which correspond to defined input parameters in the building model. These parameters are then assigned 'best-guess estimates' and 'ranges of variation' in order to generate an uncertainty-based search space. A coarse search of this space is carried out using a Monte-Carlo (MC) simulation approach to identify strong and weak parameters. This is achieved by coupling a blind Latin-Hypercube Monte-Carlo (LHMC) search with a regional sensitivity analysis (RSA). This allows the analyst to fix weak parameters and specify narrower bounds of variability for influential parameters to further refine the search space and the corresponding promising vector solutions. By adopting this multi-solution approach, predictions may be made about the effect of changes to a building (ECM analysis) while providing an associated uncertainty of these predictions. This approach is applied to three case study office buildings using the DOE-2 software and calibrating to monthly utility bills [43].

*5.4.1.3. Bayesian calibration.* It is important to consider prediction and uncertainty analysis for systems which are approximated using complex mathematical models. Bayesian calibration methods [73,74] can be used to naturally incorporate these uncertainties in the calibration process including the remaining uncertainty over the fitted parameters. These uncertainties may be propagated through the model using probabilistic sensitivity analysis [64]. Bayesian calibration methods also attempt to correct for any inadequacy of the model which is revealed by a discrepancy between the observed data and the model predictions from even the best-fitting parameter values. In addition Bayesian methods have the ability to combine multiple sources of information at varying scales and reliabilities [71].

Kennedy and O'Hagan [73] present a generic approach for the Bayesian calibration of computer models. The method is illustrated by using data from a nuclear radiation release at Tomsk-7 chemical plant, and from a more complex simulated nuclear accident exercise.

Booth et al. [71] suggest a hierarchical framework in which a top-down (macro-level) statistical model is used to infer energy consumption for (micro-level) representative individual dwellings from publically available energy consumption statistics. In this approach, a Bayesian regression method is employed for the top-down statistical model in order to account for uncertainties in the macro-level data.

#### 5.4.2. Alternative modelling techniques

*5.4.2.1. Artificial neural networks (ANN).* While not used to calibrate energy models, artificial neural networks (ANN) have been proposed as a prediction method for building energy consumption. Neto and Fiorelli [76] compared the use of EnergyPlus and artificial neural networks (ANN) in simulating energy consumption for an administration building at the University of Sao Paulo, Brazil. The

results showed that EnergyPlus consumption forecasts had an error range of ±13% for 80% of the tested 54 days of measured data. The authors concluded that the major source of uncertainties in the detailed model predictions are related to proper evaluation of lighting, equipment and occupancy. An adequate evaluation of the coefficient of performance (COP) for the unitary air conditioners serving the space also plays a very significant role in the prediction of the energy consumption of a building. The ANN models, based on simple (temperature-only input) and complex (temperature/relative humidity/solar radiation inputs) neural networks showed a fair agreement between measured and predicted energy consumption forecasts and actual values, with an average error of about 10%. While the ANN model required less manual input, it can only predict energy consumption based on past performance and therefore requires a large historic set of training data for adequate performance. Therefore, any operation changes or retrofit measures would require re-training using a new data set. Finally, the ANN model cannot provide the same insights as a detailed energy model as it is not based on physical input parameters. However, the authors conclude that there is merit in further investigating the potential for using ANN to improve methodologies for evaluation of energy consumption in air-conditioned buildings (e.g., as an substitute for complex schedule input in detailed energy models).

*5.4.2.2. PSTAR.* The primary and secondary term analysis and re-normalisation (PSTAR) method, originally proposed by Subbarao et al. [52] and later refined and extended by Burch et al. [80], and Balcomb et al. [81], utilises data from a short-term energy monitoring (STEM) test (Refer Section 5.3.1.2). In this approach, adjustments are made to major energy flows rather than to individual input parameters. This is achieved by identifying all the heat flows relevant for the building using a three stage STEM testing procedure

(1) Steady-state heat loss during constant heat input (Night).
(2) Thermal mass using cool-down test (Night).
(3) Effective solar gain by analysing change in heating/cooling load (Day).

A re-normalisation procedure (using a linear least squares method) is used to define the primary flows and subsequently compute secondary term flows thus enabling the definition of a dynamic energy balance representation of the building system. The process of data analysis and calibration from a set of defined STEM data can be automated, and is reported to yield reasonable results. However, it is dependent upon measurement accuracy. Infiltration heat loss is the major source of uncertainty and may require continuous tracer gas measurements.

*5.4.2.3. Meta-modelling.* Currently, building energy simulation models are primarily used at the building design stage, usually for the purpose of energy code compliance certification (e.g., LEED and BREEAM). As building energy models become more accurate and numerically efficient, model-based optimisation of building design and operation is becoming more practical. This model-based optimisation generally requires the combination of a whole-building energy simulation model with an optimisation tool. However, this tends to be time consuming due to the simulation and analysis time required for each model iteration. It also often leads to suboptimal results because of the detail and physical complexity of the energy model.

Eisenhower et al. [82] present an approach which aims to cut the complexity of the optimisation problem, by reducing the detailed simulation model to a simple mathematical meta-model. The method begins by sampling the parameter space of the building model around the baseline values. This is done by applying a uniform distribution and a corresponding range (±20%) of the baseline parameter value, and then using quasi-Monte Carlo (deterministic) sampling approach to provide samples within this distribution. Numerous simulations (~3000) are performed using this sample data, and an analytical meta-model is then fit to the output data. Once this process is complete, optimisation can be performed using different optimisation cost functions or optimisation algorithms with very little computational effort. Uncertainty and sensitivity analysis is also performed to identify the most influential parameters for the optimisation. A case study is explored using an EnergyPlus model of an existing building which contains over 1000 parameters. When using a cost function that penalises thermal comfort and energy, 45% annual energy reduction is achieved while simultaneously increasing thermal comfort by a factor of two.

Manfren [83] proposes an approach for calibration and uncertainty analysis in building simulation models based on the use of 'grey-box' meta-modelling techniques, combining data-driven 'black-box' models with detailed law-driven 'white-box' simulation models. This approach is applied to a real case-study office building for the verification and control of energy saving measures results. In addition, the approach is used to create a validated building simulation model for design and operational optimisation. The proposed methodology employs three models to achieve this goal (1) simple piece-wise regression model trained on real data, (2) a Gaussian process meta-model trained on computer simulation data and calibrated with respect to piece-wise regression data, and (3) a detailed simulation model directly fitted to real data. The authors propose the development of the 'black-box' Gaussian meta-model which allows performing optimisation, uncertainty and sensitivity analysis in an easier and more computationally efficient manner compared with the original 'white-box' simulation model, while maintaining comparable results. This meta-model is also used for calibrating the detailed model input variables with respect to normalised observed data (outputs). Since this approach uses computationally-efficient black-box models, it can be easily integrated with multivariate real measured data. It may also be extended to incorporate highly multivariate inputs and multiple outputs within a real-time simulation environment. The paper concludes that this approach of combining data-driven and law-driven procedures has the potential to increase the potential usefulness, transparency and applications of models for simulation-based design and optimisation of buildings.

*5.4.2.4. Simplified energy analysis procedure (SEAP).* In the early years of commercial building energy performance simulation, many solutions were quite complex and required specialists and main frame computers to run. Detailed physical models (e.g., DOE-2, EnergyPlus) also tend to be over-parameterized and can often require significant effort, experience and time to provide an accurate representation of the building. In response, simplification procedures were proposed to increase computational efficiency.

Turiel et al. [126] also proposed a simplified method of commercial building energy analysis utilising a database of previous DOE-2.1A simulations to predict the outcome of other simulations. This approach is applied to an office building with very accurate results for heating, cooling and total energy use.

Knebel [85] proposed the simplified energy analysis procedure (SEAP) in order to reduce model complexity and calibration effort. This simplification is achieved in several ways:

- The building is assumed to have only two zones (one core and one perimeter).

- Average daily data and steady-state models are used for simulation and analysis.
- One large air-handling unit (AHU) is substituted for numerous smaller ones for each zone. This is only done with similar types of AHUs.

This has been successfully applied to a number of campus and commercial buildings with great success [84,86,127,128]. This approach has since been combined with the use of signature analysis techniques (see Section 5.3.2.3) to help minimise the expertise needed to calibrate such a model [53,54].

*5.4.2.5. Systems identification.* This technique refers to the process of constructing models based only on the observed behaviour of the system (outputs) and a set of external variables (inputs), instead of constructing a detailed model based on "first principles" of well-known physical variables. Systems identification is based on work first stated in 1977 [87], but the first systematic procedure using computation tools was developed by Ljung [88]. Typically, the objective is to build a so called "black box" or "grey box" model in situations where a very detailed model would be costly and overly complex. Systems identification methods are very effective when significant amounts of data are available, as is the case with modern IT systems and advanced HVAC controls. This approach also involves an iterative procedure aimed at finding the best fit solution for model inputs.

Liu and Henze [89] applied system identification techniques to find best-tuned input settings of detailed building energy performance simulation models. This is based on a two-stage calibration process which aims to minimise the root-mean square error (RMSE) between real and simulated data. However, instead of manually adjusting the identified tuning parameters, optimisation algorithms are applied instead. Nielsen and Madsen [129] present a grey-box approach for modelling the heat consumption in district heating systems. Their approach utilises theoretical based identification of an overall model structure, followed by data-based modelling which is used to identify details of the model.

## 6. Conclusions

Buildings represent complex systems with high levels of interdependence on many external sources. The design, analysis and optimisation of modern building systems may benefit greatly from the implementation of building energy performance simulation (BEPS) tools at all stages of the building life-cycle (BLC). However, studies have found discrepancies between modelled and measured energy use in many cases where BEPS has been used to model real buildings. This undermines confidence in building simulation tools and inhibits widespread adoption.

Calibration aims to minimise discrepancies between measured and simulated data. However, due to the sheer number of inputs required for detailed building energy simulation and the limited number of measured outputs, calibration will always remain an indeterminate problem which yields a non-unique solution. Numerous approaches to model calibration have been suggested employing various combinations of analytical and/or mathematical and statistical techniques. However, no consensus has been reached on standard calibration procedures and methods that can be used generically on a wide variety of buildings. In addition, many of the current approaches to model calibration rely heavily on user knowledge, past experience, statistical expertise, engineering judgement, and an abundance of trial and error. Furthermore, when a model is established as being calibrated, the author often does not reveal the techniques used, other than stating the final result.

In summary, the issues with calibrated simulation can be broken down into seven main areas, as previously mentioned in Table 2

- **Standards**: the lack of a consensus standard on simulation calibration. There are guidelines which specify broad ranges of allowable error for building energy models. However, these are over-simplified, in that they do not account for issues such as input uncertainty/inaccuracy or the model fit to zone-level environmental data. In addition, there are no standard guidelines for model development, which leads to fragmentation of the practice of energy modelling.
- **Expense**: Due to the fragmentation of the energy modelling process, it tends to require significant effort for both model development as well as model calibration. There is no integrated standard tool-chains or file formats at present, and building data required for modelling is often unavailable. Therefore, significant expense can be incurred in building auditing, metering and model development.
- **Simplification**: One of the problems with detailed building energy simulation is the fact that they require thousands of inputs for model definition. In practice, many of these inputs are simply un-attainable or may not be practicably measureable. In addition, the data on which these models are validated is limited, generally confined to single measurements for whole building heat energy and electrical loads. Therefore, it is said that the calibration problem, as it relates to detailed models, is over-specified (i.e., too many inputs) and under-determined (i.e., too few validation points). This is a difficult problem to address, as it requires the simplification of detailed models while maintaining accuracy.
- **Inputs**: In any modelling environment, the quality of outputs are only as good as the inputs available (Garbage-in, Garbage-out). In the case of building energy modelling, the sheer number of inputs required makes it impossible to obtain accurate measurements for all parameters. In such cases, it is necessary to find ways of quantifying these parameters to a reasonable degree of accuracy without compromising model output quality;
- **Uncertainty**: Since, building energy modelling requires a degree of approximation and simplification, it is important to account for this when presenting model outputs. As shown in Section 3, there are many sources of uncertainty in building energy modelling. One of the primary sources of model uncertainty is parameter spcification uncertainty, which relates to the degree of uncertainty around each input parameter. This is often disregarded in BEPS calibration case studies, leading to questions over the accuracy of the model outputs;
- **Identification**: The calibration process, at present, can often be described as an ad-hoc procedure requiring numerous iterations of manual pragmatic user intervention based on knowledge or expert judgement. Generally, this procedure is not well defined, in that the analyst decides on model changes based on personal jusgement as opposed to quantifyable evidence. This is often difficult to define, though some studies in the literature have attempted to provide procedures for identifyaing and correcting calibration issues.
- **Automation**: With the level of manual pragmatic user intervention required during all steps of the calibration process, it is clear that any degree of automation would greatly aid this proocess. However, since many procedures require human knowledge or input, this can be difficult.

Table 6 highlights the main calibration techniques described in this review and how each provides a basis for addressing some of these issues.

**Table 6**
Addressing model calibration issues.

| Category | Classification | Approach | Calibration issues | | | | | | | References |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Standards | Expense | Simplification | Inputs | Uncertainty | Identification | Automation | |
| **Manual** | **Characterisation techniques** | AUDIT | X | | | X | | | | [40,41] |
| | | EXPERT | | | X | X | | | | [42–45] |
| | | INT | | | | X | | | | [46] |
| | | HIGH | | | | X | | | | [38,47–49] |
| | | STEM | | | | X | | | | [50–52] |
| | **Advanced graphical methods** | 3D | | | | | | X | | [47] |
| | | SIG | | | X | | | X | | [53,54] |
| | | STAT | X | | | | | | | [47] |
| | **Model simplification techniques** | BASE | | | | X | | | | [55,56] |
| | | MPE | | | X | X | | | | [57] |
| | | PARRED | | | X | X | | | | [19,20,58,59] |
| | | DISSAG | | | X | X | | | | [60–62] |
| | **Procedural extensions** | EVIDENCE | X | | | X | | X | | [21,48,59] |
| | | SA | X | | | X | | X | | [64–68] |
| | | UQ | | | | | X | | | [32,33,35,69,70] |
| **Automated** | **Optimisation techniques** | BAYES | | X | | | X | | X | [71–74] |
| | | OBJECT | X | | | | | | | [15,42,75] |
| | **Alternative modelling techniques** | ANN | | X | X | | | | X | [76–79] |
| | | PSTAR | | | X | | | | | [52,80,81] |
| | | META | | X | X | | | | | [82,83] |
| | | SEAP | | X | X | | | | | [84–86] |
| | | SYS | | | X | X | | | | [87–89] |

Based on the above extensive literature review, it is evident that the current approach to calibrating a model is at best based on an optimisation process used to identify multiple solutions within a parameter space identified from a knowledge-base of templates of influential parameters [42]. At worst, it is based on an ad-hoc approach in which the analyst manually tunes the myriad of parameters until a solution is obtained. There are no consensus standards for which approach should be used, nor is there a widespread acceptance on the validation criteria necessary for the calibration of different building energy models depending on purpose. This review aims to provide an overview of the approaches applied by practitioners to date, in attempting to overcome some of the problems associated with model calibration. While no single approach has demonstrated an ability to tackle all of the outstanding issues, it is clear that there is an extensive body of work available which can form a sound scientific basis for the development of a standardised methodology.

## 7. Graphical summary of reviewed papers

See Figs. 3–6.

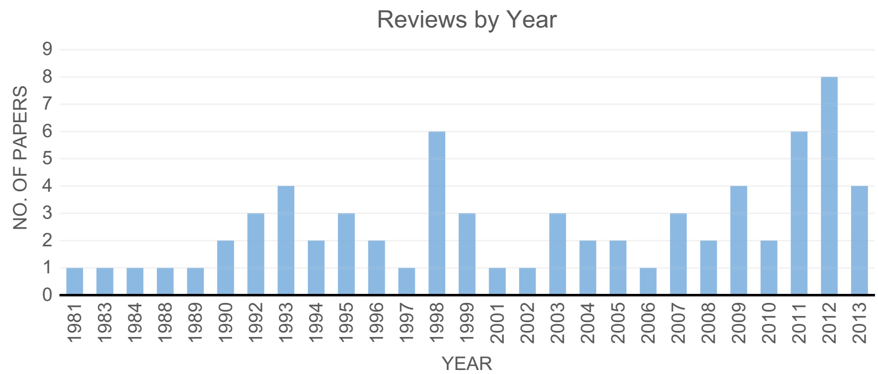



**Fig. 3.** Reviews by year.

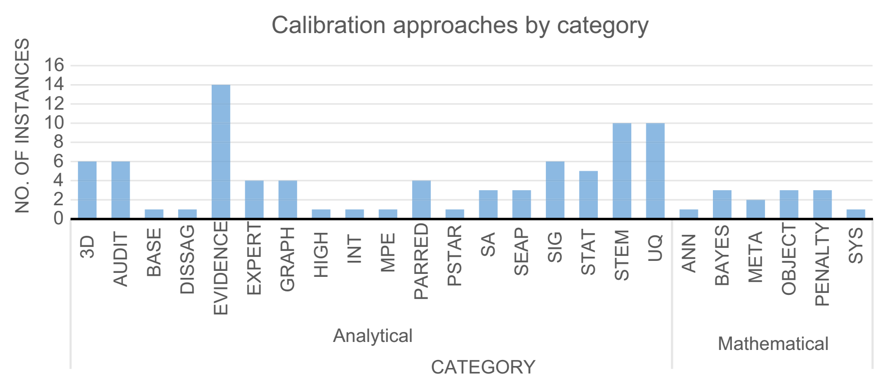



**Fig. 4.** Calibration approaches by category.

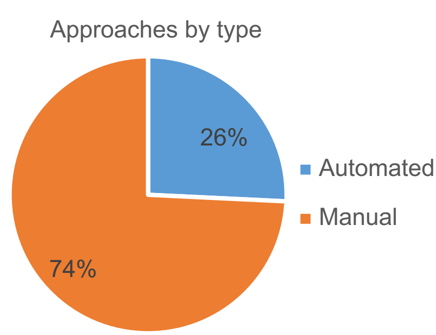



**Fig. 5.** Calibration approaches by type.

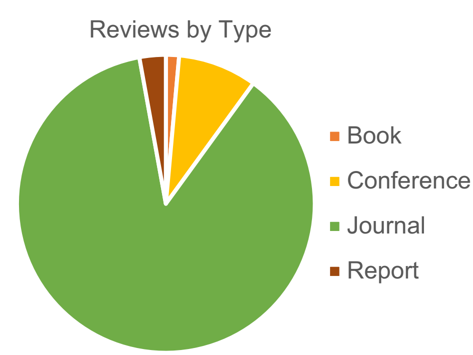



**Fig. 6.** Reviews by type.

## Acknowledgements

The authors would like to acknowledge the support of the College of Engineering and Informatics (CoEI) at National University of Ireland Galway, who funded this research.

## Appendix A

See Table A.1.

**Table A.1**
Summary of calibration papers.

| Author | Title | Journal/conference | Year | Tools | Calibration type | Analytical techniques | Math. methods | Building type | System type |
|---|---|---|---|---|---|---|---|---|---|
| Subbarao K. | Primary and secondary terms analysis and renormalisation: A unified approach to building energy simulations and short-term testing | Solar Energy Research Inst. | 1988 | PSTAR | Manual | PSTAR, STEM | N/A | Residential | N/A |
| Kaplan MB, McFerran J; Jansen J & Pratt R. | Reconciliation of a DOE-2. 1C model with monitored end-use data for a small office building | ASHRAE Transactions | 1990 | DOE-2 | Manual | STEM, PARRED | N/A | Office | Mechanical HVAC |
| Bronson DJ, Hinchey SB, Haberl JS & O'Neal DL. | A procedure for calibrating the DOE-2 simulation programme to non-weather-dependent measured loads | Proceedings Of ASHRAE Winter Meeting | 1992 | DOE-2 | Manual | 3D | N/A | Government/ Large Office | N/A |
| Waltz JP. | Practical experience in achieving high levels of accuracy in energy simulations of existing buildings | ASHRAE Transactions | 1992 | N/A | Manual | AUDIT, EXPERT, STAT | N/A | N/A | N/A |
| Katipamula S & Claridge DE. | Use of simplified system models to measure retrofit energy savings | Journal of Solar Energy Engineering | 1993 | SEAP | Manual | SEAP | N/A | Engineering Centre | DDCV with VAV |
| Carroll WL & Hitchcock RJ. | Tuning simulated building descriptions to match actual utility data: methods and implementation | ASHRAE Transactions | 1993 | DOE-2 | Automated | EXPERT | PENALTY | School | Mixed |
| Clarke JA, Strachan PA & Pernot C. | An approach to the calibration of building energy simulation models | ASHRAE Transactions | 1993 | ESP-r | Manual | HIGH, SA | N/A | Test Cell | Passive Solar (PASSYS) |
| Bou-Saada TE & Haberl JS. | An improved procedure for developing calibrated hourly simulation models | Proceedings of the 5th International IBPSA Conference | 1995 | DOE-2.1D | Automated | EVIDENCE, 3D | N/A | Office (US DOE Forrestal Complex) | Mechanical HVAC |
| Manke JM, Hittle DC & Hancock CE. | Calibrating building energy analysis models using short term test data | Journal of Solar Energy Engineering | 1996 | BLAST | Manual | STEM | N/A | School | Furnace and air-conditioning units |
| Soebarto VI. | Calibration of hourly energy simulations using hourly monitored data and monthly utility records for two case study buildings | Proceedings of the 4th International IBPSA Conference | 1997 | DOE-2 | Manual | INT | N/A | University and Municipal Building | Dual-Duct Variable Air Volume (DDVAV) |
| Haberl JS & Bou-Saada TE. | Procedures for calibrating hourly simulation models to measured building energy and environmental data | Journal of Solar Energy Engineering | 1998 | DOE-2.1D | Automated | HIGH, 3D | N/A | Office (US DOE Forrestal Complex) | Mechanical HVAC |
| Reddy TA, Deng S & Claridge DE. | Development of an inverse method to estimate overall building and ventilation parameters of large commercial buildings | Journal of Solar Energy Engineering | 1999 | DOE-2 | Manual | MPE | N/A | Commercial Office | Mechanical HVAC |
| Lunneberg TA. | Improving simulation accuracy through the use of short-term electrical end-use monitoring | Proceedings of the 6th International IBPSA Conference | 1999 | DOE-2 | Manual | STEM | N/A | Office | N/A |
| Yoon JH & Lee EJ. | Calibration procedure of energy performance simulation model for a commercial building | Proceedings from the Building Simulation Conference | 1999 | DOE-2.1E | Manual | BASE, EVIDENCE | N/A | Large Commercial Building | Package AHU |
| Liu M, ClaridgeDE, Bensouda N, Heinemeier K, Lee Seung UK & Wei G. | High Performance Commercial Building Systems Manual of Procedures for Calibrating Simulations of Building Systems | Report. California Energy Commission | 2003 | DOE-2 | Manual | SIG, STAT | N/A | Office Building/ Campus Building | SDCV/SDVAV/ DDCV/DDVAV |
| Yoon Jongho, Lee EJ & ClaridgeDE. | Calibration Procedure for Energy Performance Simulation of a Commercial Building | Journal of Solar Energy Engineering | 2003 | DOE-2.1E | Manual | BASE, EVIDENCE, STAT | N/A | Large Commercial Building | Package AHU |
| Liu S & Henze GP. | Calibration of building models for supervisory control of commercial buildings | Proceedings of the 9th International IBPSA Conference | 2005 | EnergyPlus/ GenOpt/ | Auto | N/A | SYS | N/A | AHU/VAV/ Icemaking Chiller |
| Westphal FS & Lamberts R. | Building simulation calibration using sensitivity analysis | Proceedings of the 9th International IBPSA Conference | 2005 | EnergyPlus (v1.2) | Manual | SA | N/A | Office | Mixed Mode |
| Sun Jian & Reddy T. Agami | Calibration of Building Energy Simulation Programs Using the Analytic Optimisation Approach (RP-1051) | HVAC&R Research | 2006 | DOE-2 | Auto | UQ | MC | Pseudo-Building | N/A |
| Reddy T. Agami, Maor Itzhak & Panjapornpon Chanin | Calibrating Detailed Building Energy Simulation Programs with Measured Data – Part I: General Methodology (RP-1051) | HVAC&R Research | 2007 | DOE-2 | Auto | EXPERT, UQ | MC | N/A | N/A |
| Reddy, T. Agami, Maor Itzhak & Panjapornpon Chanin | | HVAC&R Research | 2007 | DOE-2 | Auto | EXPERT, UQ | MC | Office | VAV System |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Calibrating Detailed Building Energy Simulation Programs with Measured Data – Part II: Application to Three Case Study Office Buildings (RP-1051) | | | | | | | | |
| Neto Alberto Hernandez & Fiorelli Flavio Augusto Sanzovo | Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption | Energy and Buildings | 2008 | EnergyPlus | Auto | N/A | ANN | University Building | N/A |
| Raftery Paul, Keane Marcus & Costa Andrea | Calibrating whole building energy models: Detailed case study using hourly measured data | Energy and Buildings | 2011 | | Manual | HIGH, EVIDENCE | N/A | Office | VAV Systems |
| Liu Guopeng & Liu Mingsheng | A rapid calibration procedure and case study for simplified simulation models of commonly used HVAC systems | Building and Environment | 2011 | SEAP | Manual | SEAP, SIG | N/A | Office | VAV systems, CV system |
| Raftery Paul, Keane Marcus & O'Donnell James | Calibrating whole building energy models: An evidence-based methodology | Energy and Buildings | 2011 | EnergyPlus | Manual | EVIDENCE, PARRED | N/A | N/A | N/A |
| Heo Y, Choudhary R & Augenbroe GA. | Calibration of building energy models for retrofit analysis under uncertainty | Energy and Buildings | 2012 | EnergyPlus | Auto | UQ | BAYES | Office | Gas Condensing Boiler/Natural Ventilation |
| Coakley Daniel, Raftery Paul & Molloy Padraig | Calibration of Whole Building Energy Simulation Models: Detailed Case Study of a Naturally Ventilated Building Using Hourly Measured Data | Proceedings of Building Simulation & Optimisation Conference | 2012 | EnergyPlus | Auto | HIGH, EVIDENCE, STEM | PEN | University Building | Natural Ventilation |
| Manfren Massimiliano, Aste Niccola & Moshksar Reza | Calibration and uncertainty analysis for computer models: "A meta-model based approach for integrated building energy simulation" | Applied Energy | 2013 | N/A | Auto | UQ | BAYES, META | Office Building/ Campus Building | N/A |
| Booth AT, Choudhary R, & Spiegelhalter DJ. | A hierarchical Bayesian framework for calibrating micro-level models with macro-level data | Journal of Building Performance Simulation | 2013 | N/A | Auto | UQ | BAYES | Residential | Varies |
| O'Neill Z & Eisenhower B. | Leveraging the analysis of parametric uncertainty for building energy model calibration | Building Simulation | 2013 | EnergyPlus 6.0 | Auto | SA, UQ | META | Military Barracks | VAV System |

N/A: Not applicable to this case; X: information/paper not available at the time of writing.

## References


[1] Saltelli, Andrea, Ratto Marco, Andres Terry, Campolongo Francesca, Cariboni Jessica, Gatelli Debora, et al. Global sensitivity analysis: the primer. Chichester, West Sussex PO19 8SQ, England: John Wiley & Sons Ltd; 2008. ISBN: 978-0-470-05997-5.

[2] Young P, Parkinson S, Lees M. Simplicity out of complexity in environmental modelling: Occam's razor revisited. J Appl Stat 1996;23:165–210.

[3] Florita AR, Henze GP. Whole building fault diagnostics – a Bayesian approach. Intell Build Oper Work, Boulder, CO 2013. In: procceding of 2013 intelligent building operations workshop, June 20–22, 2013, Boulder, CO, USA, (see 〈http://www.ibpsa.us/IBO2013.shtml〉).

[4] Reddy TA, Andersen KK. An evaluation of classical steady-state off-line linear parameter estimation methods applied to chiller performance data. HVAC&R Res 2002;8:101–24.

[5] Winkelmann F, Birdsall B, Buhl W. DOE-2 supplement: version 2.1 e. LBL-34947. Berkeley: Calif Lawrence Berkeley Lab; 1993.

[6] Crawley D, Lawrie L. EnergyPlus: creating a new-generation building energy simulation program. Energy Build 2001;33:319–31.

[7] Klein SA, Beckman WA, Mitchell JW, Duffie JA, Duffie NA, Freeman TL, et al. TRNSYS 16-A, transient system simulation program. Madison, WI, USA: University of Wisconsin; 1979. For more info, 〈http://sel.me.wisc.edu/trnsys/〉.

[8] ESRU. ESP-r. Glasgow, Scotland G1 1XJ: Energy Systems Research Unit, University of Strathclyde; 1974.

[9] Crawley DBD, Hand JWJ, Kummert M, Griffith BT. Contrasting the capabilities of building energy performance simulation programs. Build Environ 2008;43:661–73.

[10] Augenbroe G. Trends in building simulation. Build Environ 2002;37:891–902.

[11] Tupper K, Franconi E, Chan C, Fluhrer C, Jenkins M, Hodgin S. Pre-Read for BEM innovation summit. building energy model. In: Innovation summit. Boulder, CO; 2011.

[12] Karlsson F, Rohdin P, Persson ML. Measured and predicted energy demand of a low energy building: important aspects when using building energy simulation. Build Serv Eng Res Technol 2007;28:223–35.

[13] Turner C, Frankel M. Energy performance of LEED for new construction buildings. Washington, DC: New Buildings Institute; 2008.

[14] Scofield JH. Do LEED-certified buildings save energy? Not really… Energy Build 2009;41:1386–90.

[15] Carroll WL, Hitchcock RJ. Tuning simulated building descriptions to match actual utility data: methods and implementation. ASHRAE Trans 1993;99:928–34.

[16] Kaplan M, Jones B, Jansen J. DOE-2.1 C model calibration with monitored end-use data. In: Proc from ACEEE 1990 summer study energy efficient building: vol. 10; 1990. p. 115–25.

[17] Troncoso R. A hybrid monitoring-modeling procedure for analyzing the performance of large central chilling plants. In: Proceedings of the 6th international IBPSA conference, vol. 97. Prague, Czech Republic: International Building Performance Simulation Association; 1997. p. 421–8.

[18] Diamond S, Hunn B. Comparison of DOE-2 computer program simulations to metered data for seven commercial buildings. ASHRAE Trans 1981;87:1222–31.

[19] Kaplan M, McFerran J, Jansen J, Pratt R. Reconciliation of a DOE2. 1C model with monitored end-use data for a small office building. ASHRAE Trans 1990;96:982–93.

[20] Bronson DJ, Hinchey SB, Haberl JS, O'Neal DL. Procedure for calibrating the DOE-2 simulation program to non-weather-dependent measured loads. In: Proceedings of the ASHRAE winter meeting. Anaheim, CA, USA; 1992. p. 636–52.

[21] Bou-Saada TE, Haberl JS. An improved procedure for developing calibrated hourly simulation models. In: Proceedings of the 5th international IBPSA conference. Montréal, Canada; 1995.

[22] Kreider J, Haberl J. Predicting hourly building energy usage. ASHRAE J 1994;36:72–81.

[23] Kreider JF, Haberl JS. Predicting hourly building energy use: the great energy predictor shootout: overview and discussion of results. ASHRAE Trans 1994;100:1104–18.

[24] ASHRAE. Guideline 14-2002: measurement of energy and demand savings. Atlanta, GA 30329: American Society of Heating, Refrigerating and Air-Conditioning Engineers; 2002.

[25] EVO. International performance measurement & verification protocol. Efficiency Valuation Organisation; 2007.

[26] US DOE. M&V guidelines: measurement and verification for federal energy projects version 3.0. US Department Of Energy; 2008.

[27] Funtowicz Silvio O, Ravetz Jerome R. Uncertainty and quality in science for policy, vol. 15. Norwell, MA 02061, U.S.A: Kluwer Academic Publishers; 1990. ISBN: 0-7923-0799-2.

[28] Rosen Robert. Life itself: a comprehensive inquiry into the nature, origin, and fabrication of life. New York, NY: Columbia University Press; 1991. ISBN: 978-0231075657.

[29] Beven K. Prophecy, reality and uncertainty in distributed hydrological modelling. Adv Water Resour 1993;16:41–51.

[30] Aronica G, Hankin B, Beven K. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. Adv Water Resour 1998;22:349–65.

[31] Oreskes, Naomi, Kristin Shrader-Frechette, Kenneth Belitz. Verification, validation, and confirmation of numerical models in the earth sciences. Science 1994;263(5147):641–6.

[32] De Wit S, Augenbroe G. Analysis of uncertainty in building design evaluations and its implications. Energy Build 2002;34:951–8.
[33] Macdonald IA. Quantifying the effects of uncertainty in building simulation [Ph.D. thesis]. University of Strathclyde; 2002.
[34] Macdonald I, Strachan P. Practical application of uncertainty analysis. Energy Build 2001;33:219–27.
[35] Macdonald I, Clarke JA, Strachan PA. Assessing uncertainty in building simulation. In: Proceedings of the 6th international IBPSA conference. Kyoto, Japan: International Building Performance Simulation Association; 1999.
[36] Moon H. Assessing mold risks in buildings under uncertainty [Ph.D. thesis]. GA, USA: Georgia Institute of Technology; 2005.
[37] Maile T, Fischer M. Formalizing approximations, assumptions, and simplifications to document limitations in building energy performance simulation. Comparing measured and simulated building energy performance data; 2010.
[38] Clarke JA, Strachan PA, Pernot C. An approach to the calibration of building energy simulation models. ASHRAE Trans 1993;99:917.
[39] Reddy TA. Literature review on calibration of building energy simulation programs: uses, problems, procedures, uncertainty, and tools. ASHRAE Trans 2006;112:226–40.
[40] Shapiro I. Energy audits in large commercial office buildings. ASHRAE J 2009;51:18–27.
[41] Waltz J. Practical experience in achieving high levels of accuracy in energy simulations of existing buildings. ASHRAE Trans. 1992;98(part 1):606–17 paper number AN-92-1-2.
[42] Reddy TA, Maor I, Panjapornpon C. Calibrating detailed building energy simulation programs with measured data – Part I: general methodology (RP-1051). HVAC&R Res 2007;13:221–41.
[43] Reddy TA, Maor I, Panjapornpon C. Calibrating detailed building energy simulation programs with measured data – Part II: application to three case study office buildings (RP-1051). HVAC&R Res 2007;13:243–65.
[44] Lebot B. Investigation of an expert system to calibrate a computer simulation program; 1987.
[45] Hitchcock RJ, Carroll WL, Birdsall BE. RESEM: retrofit energy savings estimation model reference manual. Berkeley, CA 94720: Applied Science Division, Lawrence Berkeley Laboratory (LBL), University of California; 1991.
[46] Soebarto V. Calibration of hourly energy simulations using hourly monitored data and monthly utility records for two case study buildings. In: Proceedings of the 4th international IBPSA conference. Prague, Czech Republic; 1997.
[47] Haberl JSJ, Bou-Saada TE. Procedures for calibrating hourly simulation models to measured building energy and environmental data. J Sol Energy Eng 1998;120:193–204.
[48] Raftery P, Keane M, Costa A. Calibrating whole building energy models: detailed case study using hourly measured data. Energy Build 2011;43:3666–79.
[49] Coakley D, Raftery P, Molloy P. Calibration of whole building energy simulation models: detailed case study of a naturally ventilated building using hourly measured data. Build Simul Optim 2012;57–64. In: Conference proceedings - Wright Jonathan, Cook Malcolm, (editors.): 2012, Proceedings of the 2012 building simulation and optimization conference. 10–11 September 2012, Loughborough, Leicestershire, UK; ISBN: 978-1-897911-42-6.
[50] Lunneberg T. Improving simulation accuracy through the use of short-term electrical end-use monitoring. In: Proceedings of the 6th international IBPSA conference. Kyoto, Japan; 1999. p. 1–8.
[51] Manke J, Hittle D, Hancock C. Calibrating building energy analysis models using short term test data. In: Proceedings of the 1996 ASME international solar engineering conference. San Antonio, TX; 1996. p. 369–78.
[52] Subbarao K. Primary and secondary terms analysis and renormalization: a unified approach to building energy simulations and short-term testing. Golden, CO; 1988.
[53] Liu M, Claridge DE, Bensouda N, Heinemeier K, Lee SU, Wei G. High performance commercial building systems manual of procedures for calibrating simulations of building systems. California, USA: California Energy Commission (CEC); 2003.
[54] Liu G, Liu M. A rapid calibration procedure and case study for simplified simulation models of commonly used HVAC systems. Build Environ 2011;46:409–20.
[55] Yoon J, Lee EJ, Claridge DE. Calibration procedure for energy performance simulation of a commercial building. J Sol Energy Eng 2003;125:251.
[56] Yoon J, Lee E. Calibration procedure of energy performance simulation model for a commercial building. In: Proceedings of building simulations 1999. Kyoto, Japan: International Building Performance Simulation Association; 1999.
[57] Reddy T, Deng S, Claridge D. Development of an inverse method to estimate overall building and ventilation parameters of large commercial buildings. J Sol Energy Eng 1999;121:40–6.
[58] Hadley D. Daily variations in HVAC system electrical energy consumption in response to different weather conditions. Energy Build 1993;19:235–47.
[59] Raftery P, Keane M, O'Donnell J, O'Donnell J. Calibrating whole building energy models: an evidence-based methodology. Energy Build 2011;43:2356–64.
[60] Akbari H. Validation of an algorithm to disaggregate whole-building hourly electrical load into end uses. Energy 1995;20:1291–301.
[61] Akbari H, Konopacki S. Application of an end-use disaggregation algorithm for obtaining building energy-use data. J Sol Energy Eng 1998;120:205–10.
[62] Akbari H, Heinemeier K, Flora D, LeConiac P. Analysis of commercial whole-building 15-minute-interval electric load data. ASHRAE Trans 1988;94(2) OT-88-01-1. [Accessed at: 〈https://publications.lbl.gov/islandora/object/ir%3A113489〉].
[63] Coakley D, Raftery P, Molloy P, White G. Calibration of a detailed BES model to measured data using an evidence-based analytical optimisation approach. In: Proceedings of the 12th international IBPSA conference. Sydney, Australia; 2011.
[64] Lomas K, Eppel H. Sensitivity analysis techniques for building thermal simulation programs. Energy Build 1992;19:1–44.
[65] Saltelli A, Tarantola S, Campolongo F, Ratto M. Sensitivity analysis in practice: a guide to assessing scientific models. Hoboken, NJ: Wiley; 2004.
[66] Lam JC, Hui SCM. Sensitivity analysis of energy performance of office buildings. Build Environ 1996;31:27–39.
[67] Eisenhower B, O'Neill Z. Uncertainty and sensitivity decomposition of building energy models. J Build Perform Simul 2012;5:37–41.
[68] Westphal F, Lamberts R. Building simulation calibration using sensitivity analysis. In: Proceedings of the 9th international IBPSA conference. Montréal, Canada; 2005. p. 1331–8.
[69] O'Neill Z, Eisenhower B, Fonoberov V, Bailey T. Calibration of a building energy model considering parametric uncertainty. ASHRAE Trans 2012, 189–196;118.
[70] Heo Y, Choudhary R, Augenbroe G. Calibration of building energy models for retrofit analysis under uncertainty. Energy Build 2012;47:550–60.
[71] Booth T, Choudhary R, Spiegelhalter DJ. A hierarchical Bayesian framework for calibrating micro-level models with macro-level data. J Build Perform Simul 2013;6:1–26.
[72] O'Hagan a. Bayesian analysis of computer code outputs: a tutorial. Reliab Eng Syst Saf 2006;91:1290–300.
[73] Kennedy MC, O'Hagan A. Bayesian calibration of computer models. J R Stat Soc Ser B (Stat Methodol) 2001;63:425–64.
[74] MacKay D. Bayesian nonlinear modeling for the prediction competition. ASHRAE Trans 1994;100:1053–62.
[75] Sun J, Reddy TA. Calibration of building energy simulation programs using the analytic optimization approach (RP-1051). HVAC&R Res 2006;12:177–96.
[76] Neto AH, Fiorelli FAS. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. Energy Build 2008;40:2169–76.
[77] Karatasou S, Santamouris M, Geros V. Modeling and predicting building's energy use with artificial neural networks: methods and results. Energy Build 2006;38:949–58.
[78] Kalogirou SA, Bojic M. Artificial neural networks for the prediction of the energy consumption of a passive solar building. Energy 2000;25:479–91.
[79] Mihalakakou G, Santamouris M, Tsangrassoulis A. On the energy consumption in residential buildings. Energy Build 2002;34:727–36.
[80] Burch J, Subbarao K, Lekov A, Warren M, Norford L. Short-term energy monitoring in a large commercial building. ASHRAE Trans 1990;96:1459–77.
[81] Balcomb J, Burch J, Subbarao K. Short-term energy monitoring of residences. ASHRAE Trans 1993;99:935–44.
[82] Eisenhower B, O'Neill Z, Narayanan S, Fonoberov Va, Mezić I. A methodology for meta-model based optimization in building energy models. Energy Build 2012;47:292–301.
[83] Manfren M, Aste N, Moshksar R. Calibration and uncertainty analysis for computer models – a meta-model based approach for integrated building energy simulation. Appl Energy 2013;103:627–41.
[84] Katipamula S, Claridge D. Use of simplified system models to measure retrofit energy savings. J Sol Energy Eng 1993;115:57–68.
[85] Knebel D. Simiplified energy analysis using the modified bin method. Atlanta, GA, USA: American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE); 1983.
[86] Liu M, Claridge DE. Use of calibrated HVAC system models to optimize system operation. J Sol Energy Eng 1998;120:131–8.
[87] Goodwin G, Payne R. Dynamic system identification: experiment design and data analysis. San Diego, California: Academic Press, Inc; 1977.
[88] Ljung L. System identification: theory for the user. Englewood Cliffs, New Jersey 07632: Prentice-Hall; 1987.
[89] Liu S, Henze GP. Calibration of building models for supervisory control of commercial buildings. In: Proceedings of the 9th international building performance simulation association conference. Montréal, Canada; 2005.
[90] Thumann A, Younger W. Handbook of energy audits. 7th ed.. Lilburn, GA: Fairmont Press; 2008.
[91] ASHRAE. Procedures for commercial building energy audits. 2nd ed. Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers; 2011.
[92] Lyberg M. Source book for energy auditors, vols. 1–2. Stockholm, Sweden: Swedish Council for Building Research; 1987.
[93] CEC. Guide to preparing feasibility studies of energy efficiency projects. California, USA: California Energy Commission (CEC); 2000.
[94] Ganji AR, Gilleland B. Investment grade energy audit. In: Proceedings of the 25th world energy engineering congress. Atlanta, GA; 2002, p. 1–5.
[95] CASCADE Consortium. Deliverable D2.1 CASCADE methodology for energy efficient airports; 2012.
[96] Adnot J. Field benchmarking and market development for audit methods in air conditioning (AUDITAC) final report. Intell Energy Eur 2007 [Project Report] field benchmarking and market development for Audit methods in air conditioning (AUDITAC) final report, Intelligent Energy Europe; February 2007. 105 pp. Available at: 〈http://www.eaci-projects.eu/iee/page/Page.jsp?op=project_detail&prid=1439〉.

[97] Standards Australia. AS/NZ 3598:2000 Australian/New Zealand standard: energy audits. Australia/New Zealand; 2000.
[98] ASHRAE. RP-351-energy audit input procedures and forms. Atlanta, GA; 1983.
[99] TRC. DOE-2: comparison with measured data: design and operational energy studies in a new high-rise office building. New York, USA; 1984.
[100] Norford LK, Socolow RH, Hsieh ES, Spadaro GV. Two-to-one discrepancy between measured and predicted performance of a "low-energy" office building: insights from a reconciliation based on the DOE-2 model. Energy Build 1994;21:121–31.
[101] Hunn B, Banks J, Reddy S. Energy analysis of the Texas capitol restoration. Proceedings of the 8th symposium on improving building systems in. hot and humid climate. Dallas, TX; 1992, p. 165–73.
[102] Mccray JA, Bailey PL, Parker JL, Gillman R. Using data visualization tools for the calibration of hourly DOE-2. 1 simulations calibration of simulation models. In: Proceedings of the 3rd international IBPSA conference. Madison, WI; 1995. p. 461–6.
[103] Haberl JS, Bronson DJ, Hinchey SB, O'Neal DL. Graphical tools to help calibrate the DOE-2 simulation program to non-weather dependent measured loads. ASHRAE J 1993;35:27–32.
[104] Haberl J, Abbas M. Development of graphical indices for viewing building energy data: Part II. J Sol Energy Eng 1998;120:162–7.
[105] Haberl J, Abbas M. Development of graphical indices for viewing building energy data: Part I. J Sol Energy Eng 1998;120:156–61.
[106] Haberl J, Sparks R, Culp C. Exploring new techniques for displaying complex building energy consumption data. Energy Build 1996;24:27–38.
[107] Christensen C. Digital and color energy maps for graphic display of hourly data. In: Proceedings of the 9th national passive solar conference, vol. 1. Columbus, Ohio; 1984.
[108] Raftery P, Keane M. Visualising patterns in building performance data. In: Proceedings of the 12th international IBPSA conference. Sydney, Australia; 2011.
[109] Abbas M. Development of graphical indices for building energy data [M.S. thesis]. Texas A&M University; 1993.
[110] Wilde P, Tian W. Identification of key factors for uncertainty in the prediction of the thermal performance of an office building under climate change. Build Simul 2009;2:157–74.
[111] Wang L, Mathew P, Pang X. Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building. Energy Build 2012;53:152–8.
[112] Wei G, Liu M, Claridge D. Signatures of heating and cooling energy consumption for typical AHUs. In: Proceedings of the eleventh symposium on improving building systems in hot and humid climates. Forth Worth, TX; 1998. p. 387–402.
[113] Yu FW, Chan KT. Energy signatures for assessing the energy performance of chillers. Energy Build 2005;37:739–46.
[114] Liu M, Wei G, Claridge D. Calibrating AHU models using whole building cooling and heating energy consumption data. In: Proceedings of the ACEEE 1998 summer study energy efficient building: vol. 3; 1998. p. 229–41.
[115] Parker J, Cropper P, Shao L. A calibrated whole building simulation approach to assessing retrofit options for Birmingham Airport. In: Proceedings of the building simulation and optimization conference. Loughborough, UK; 2012. p. 49–56.
[116] Monfet D, Charneux R, Zmeureanu R, Lemire N. Calibration of a building energy model using measured data. ASHRAE Trans 2009;115:348–59.
[117] Saltelli A. Sensitivity analysis for importance assessment. Risk Anal 2002;22:579–90.
[118] Saltelli A. Global sensitivity analysis: an introduction. Ispra, Italy: European Commission, Joint Research Centre; 2005.
[119] Henninger RH, Witte MJ, Crawley DB. Analytical and comparative testing of EnergyPlus using IEA HVAC BESTEST E100-E200 test suite. Energy Build 2004;36:855–63.
[120] Ryan EM, Sanquist TF. Validation of building energy modeling tools under idealized and realistic conditions. Energy Build 2012;47:375–82.
[121] Neymark J, Judkoff R, Knabe G, Le H-T, Dürig M, Glass A. Applying the building energy simulation test (BESTEST) diagnostic method to verification of space conditioning equipment models used in whole-building energy simulation programs. Energy Build 2002;34:917–31.
[122] Lomas KJ, Eppel H, Martin CJ, Bloomfield DP. Empirical validation of building energy simulation programs. Energy Build 1997;26:253–75.
[123] Pal Vineeta, William L Carroll, Bourassa Norman. RESEM-CA: Validation and testing, LBNL No. 52003 (HPBCS E2P2.2T2b). California, USA: Ernest Orlando Lawrence Berkeley National Laboratory (LBNL); 2002 [see 〈http://eetd.lbl.gov/sites/all/files/publications/lbnl-52003.pdf〉].
[124] Carroll WL, Birdsall BE, Hitchcock RJ, Kammerud RC. RESEM: an evaluation tool for energy retrofits in institutional buildings. In: Proceedings of the 1st international IBPSA conference. Vancouver, Canada; 1989.
[125] Lavigne K. Assisted calibration in buildings simulation – algorithm description and case studies. In: Proceedings of the 11th international IBPSA conference. Glasgow, Scotland; 2009.
[126] Turiel I, Boschen R, Seedall M, Levine M. Simplified energy analysis methodology for commercial buildings. Energy Build 1984;6:67–83.
[127] Liu Mingsheng, David E. Claridge, Application of calibrated HVAC system models to identify component malfunctions and to optimize operation and control schedules. In: Solar engineering 1995–proceedings of the 1995, ASME/JSME/JSES international solar energy conference, vol. 1. Maui, Hawaii; March 19-24, 1995. pp. 209–217.
[128] Liu M, Song L, Wei G, Claridge DE. Simplified building and air-handling unit model calibration and applications. J Sol Energy Eng 2004;126:601.
[129] Nielsen HA, Madsen H. Modelling the heat consumption in district heating systems using a grey-box approach. Energy Build 2006;38:63–71.